

What is claimed is:

1. A method for identifying a subject at risk of breast cancer, which comprises detecting the presence or absence of one or more polymorphic variations associated with breast cancer in a nucleic acid sample from a subject, wherein the one or more polymorphic variations are detected in a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (d) a fragment of a nucleotide sequence of (a), (b), or (c);

whereby the presence of the polymorphic variation is indicative of the subject being at risk of breast cancer.

2. The method of claim 1, which further comprises obtaining the nucleic acid sample from the subject.

3. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in SEQ ID NO: 1 is selected from the group consisting of 50479, 50567, 51865, 58349, 58464, 68551, 70419, 73014, 75242, 75505, 78067, 78265, 78290, 78956, 79196, 81886, 82399, 83380, 86645, 86704, 88134, 96416, 96429, 96535, 110920, 14117, 117297, 117617, 118870, 123740, 123792, 123876, 123880, 126685, 126696, 126990, 127939, 146146, and 146312.

4. The method of claim 1, wherein the one or more polymorphic variations are detected within a region spanning positions 78265 to 123880, spanning positions 78265 to 96535, or spanning positions 117297 to 123880 in SEQ ID NO: 1.

5. The method of claim 3, wherein a polymorphic variation is detected at position 78265 in SEQ ID NO: 1.

6. The method of claim 3, wherein a polymorphic variation is detected at position 96535 in SEQ ID NO: 1.

7. The method of claim 3, wherein a polymorphic variation is detected at position 117617 in SEQ ID NO: 1.

8. The method of claim 1, wherein the one or more polymorphic variations are detected at one or more positions in linkage disequilibrium with one or more positions in SEQ ID NO: 1 selected from the group consisting of 50479, 50567, 51865, 58349, 58464, 68551, 70419, 73014, 75242, 75505, 78067, 78265, 78290, 78956, 79196, 81886, 82399, 83380, 86645, 86704, 88134, 96416, 96429, 96535, 110920, 14117, 117297, 117617, 118870, 123740, 123792, 123876, 123880, 126685, 126696, 126990, 127939, 146146, and 146312.

9. The method of claim 1, wherein detecting the presence or absence of the one or more polymorphic variations comprises:
hybridizing an oligonucleotide to the nucleic acid sample, wherein the oligonucleotide is complementary to a nucleotide sequence in the nucleic acid and hybridizes to a region adjacent to the polymorphic variation;
extending the oligonucleotide in the presence of one or more nucleotides, yielding extension products; and
detecting the presence or absence of a polymorphic variation in the extension products.

10. The method of claim 1, wherein the subject is a human.

11. A method for identifying a polymorphic variation associated with breast cancer proximal to an incident polymorphic variation associated with breast cancer, which comprises:
identifying a polymorphic variation proximal to the incident polymorphic variation associated with breast cancer, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:
(a) a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising the polymorphic variation;
determining the presence or absence of an association of the proximal polymorphic variant with breast cancer.

12. The method of claim 11, wherein the incident polymorphic variation is at a position in SEQ ID NO: 1 selected from the group consisting of 50479, 50567, 51865, 58349, 58464, 68551, 70419, 73014, 75242, 75505, 78067, 78265, 78290, 78956, 79196, 81886, 82399, 83380, 86645, 86704, 88134, 96416, 96429, 96535, 110920, 14117, 117297, 117617, 118870, 123740, 123792, 123876, 123880, 126685, 126696, 126990, 127939, 146146, and 146312.

13. The method of claim 1, wherein the incident polymorphic variation is within a region spanning positions 78265 to 123880, spanning positions 78265 to 96535, or spanning positions 117297 to 123880 in SEQ ID NO: 1.

14. The method of claim 11, wherein the proximal polymorphic variation is within a region between about 5 kb 5' of the incident polymorphic variation and about 5 kb 3' of the incident polymorphic variation.

15. The method of claim 11, which further comprises determining whether the proximal polymorphic variation is in linkage disequilibrium with the incident polymorphic variation.

16. The method of claim 11, which further comprises identifying a second polymorphic variation proximal to the identified proximal polymorphic variation associated with breast cancer and determining if the second proximal polymorphic variation is associated with breast cancer.

17. The method of claim 16, wherein the second proximal polymorphic variant is within a region between about 5 kb 5' of the incident polymorphic variation and about 5 kb 3' of the proximal polymorphic variation associated with breast cancer.

18. An isolated nucleic acid comprising a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (d) a fragment of a nucleotide sequence of (a), (b), or (c); and
- (e) a nucleotide sequence complementary to the nucleotide sequences of (a), (b), (c), or (d);

wherein the nucleotide sequence comprises a nucleotide at a position in Figures 1A-1YYYY associated with breast cancer selected from the group consisting of a guanine at position 78265, a guanine at position 79169, a thymine at position 96429, a cytosine at position 96535, a thymine at position 117297, a cytosine at position 117617, a guanine at position 118870, and a cytosine at position 123880.

19. An oligonucleotide comprising a nucleotide sequence complementary to a portion of the nucleotide sequence of (a), (b), (c), or (d) in claim 18, wherein the 3' end of the oligonucleotide is adjacent to a polymorphic variation associated with breast cancer.

20. A microarray comprising an isolated nucleic acid of claim 18 linked to a solid support.

21. An isolated polypeptide encoded by the isolated nucleic acid sequence of claim 18.

22. A method for identifying a candidate molecule that modulates cell proliferation, which comprises:

(a) introducing a test molecule to a system which comprises a nucleic acid comprising a nucleotide sequence selected from the group consisting of:

- (i) a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (ii) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (iii) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;

(iv) a fragment of a nucleotide sequence of (i), (ii), or (iii); or
introducing a test molecule to a system which comprises a protein encoded by a nucleotide sequence of (i), (ii), (iii), or (iv); and

(b) determining the presence or absence of an interaction between the test molecule and the nucleic acid or protein,

whereby the presence of an interaction between the test molecule and the nucleic acid or protein identifies the test molecule as a candidate molecule that modulates cell proliferation.

23. The method of claim 22, wherein the system is an animal.

24. The method of claim 22, wherein the system is a cell.

25. The method of claim 22, wherein the nucleotide sequence comprises one or more polymorphic variations associated with breast cancer.

26. The method of claim 25, wherein the one or more polymorphic variations associated with breast cancer are at one or more positions in SEQ-ID NO: 1 is selected from the group consisting of 50479, 50567, 51865, 58349, 58464, 68551, 70419, 73014, 75242, 75505, 78067, 78265, 78290, 78956, 79196, 81886, 82399, 83380, 86645, 86704, 88134, 96416, 96429, 96535, 110920, 14117, 117297, 117617, 118870, 123740, 123792, 123876, 123880, 126685, 126696, 126990, 127939, 146146, and 146312.

27. The method of claim 25, wherein the one or more polymorphic variations associated with breast cancer are in a region spanning positions 78265 to 123880, spanning positions 78265 to 96535, or spanning positions 117297 to 123880 in SEQ ID NO: 1.

28. A method for treating breast cancer in a subject, which comprises administering a candidate molecule identified by the method of claim 23 to a subject in need thereof, whereby the candidate molecule treats breast cancer in the subject.

29. A method for identifying a candidate therapeutic for treating breast cancer, which comprises:

- (a) introducing a test molecule to a system which comprises a nucleic acid comprising a nucleotide sequence selected from the group consisting of:
 - (i) a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
 - (ii) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
 - (iii) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
 - (iv) a fragment of a nucleotide sequence of (i), (ii), or (iii); orintroducing a test molecule to a system which comprises a protein encoded by a nucleotide sequence of (i), (ii), (iii), or (iv); and
 - (b) determining the presence or absence of an interaction between the test molecule and the nucleic acid or protein,
- whereby the presence of an interaction between the test molecule and the nucleic acid or protein identifies the test molecule as a candidate therapeutic for treating breast cancer.

30. A method for treating breast cancer in a subject, which comprises contacting one or more cells of a subject in need thereof with a nucleic acid, wherein the nucleic acid comprises a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (d) a fragment of a nucleotide sequence of (a), (b), or (c); and
- (e) a nucleotide sequence complementary to the nucleotide sequences of (a), (b), (c), or (d);

whereby contacting the one or more cells of the subject with the nucleic acid treats breast cancer in the subject.

31. The method of claim 30, wherein the nucleic acid is RNA or PNA.

32. The method of claim 31, wherein the nucleic acid is duplex RNA.

33. The method of claim 32, wherein a strand of the duplex RNA comprises a nucleotide sequence selected from the group consisting of AAGGGUCUCCAAACGUCCACA (SEQ ID NO:); AAGUAUCACAUUCACAGGAUG (SEQ ID NO:) and AAUGCCACAGUACACCAGUCG (SEQ ID NO:).

34. A method for treating breast cancer in a subject, which comprises contacting one or more cells of a subject in need thereof with a protein, wherein the protein is encoded by a nucleotide sequence which comprises a polynucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (d) a fragment of a nucleotide sequence of (a), (b), or (c);

whereby contacting the one or more cells of the subject with the protein treats breast cancer in the subject.

35. A method for treating breast cancer in a subject, which comprises:
detecting the presence or absence of one or more polymorphic variations associated with breast cancer in a nucleic acid sample from a subject, wherein the one or more polymorphic variation are detected in a nucleotide sequence selected from the group consisting of:
(a) a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
(b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
(c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
(d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising the polymorphic variation; and
administering a breast cancer treatment to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.
36. The method of claim 35, wherein the one or more polymorphic variations are at one or more positions in SEQ ID NO: 1 selected from the group consisting of 50479, 50567, 51865, 58349, 58464, 68551, 70419, 73014, 75242, 75505, 78067, 78265, 78290, 78956, 79196, 81886, 82399, 83380, 86645, 86704, 88134, 96416, 96429, 96535, 110920, 14117, 117297, 117617, 118870, 123740, 123792, 123876, 123880, 126685, 126696, 126990, 127939, 146146, and 146312.
37. The method of claim 36, wherein the one or more polymorphic variations are in a region spanning positions 78265 to 123880, spanning positions 78265 to 96535, or spanning positions 117297 to 123880 in SEQ ID NO: 1.
38. The method of claim 35, which further comprises extracting and analyzing a tissue biopsy sample from the subject.
39. The method of claim 35, wherein the treatment is chemotherapy, surgery, radiation therapy, and combinations of the foregoing.
40. The method of claim 39, wherein the chemotherapy is selected from the group consisting of cyclophosphamide (Cytoxan), methotrexate (Amethopterin, Mexate, Folex), fluorouracil (Fluorouracil, 5-Fu, Adrucil), cyclophosphamide, doxorubicin (Adriamycin), and combinations of the foregoing.

41. The method of claim 40, wherein the combinations are selected from the group consisting of cyclophosphamide (Cytosan), methotrexate (Amethopterin, Mexate, Folex), and fluorouracil (Fluorouracil, 5-Fu, Adrucil); cyclophosphamide, doxorubicin (Adriamycin), and fluorouracil; and doxorubicin and cyclophosphamide.

42. A method for detecting or preventing breast cancer in a subject, which comprises:
detecting the presence or absence of one or more polymorphic variations associated with breast cancer in a nucleic acid sample from a subject, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising the polymorphic variation; and

administering a breast cancer preventative or detection procedure to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

43. The method of claim 42, wherein the one or more polymorphic variations are at one or more positions in SEQ ID NO: 1 is selected from the group consisting of 50479, 50567, 51865, 58349, 58464, 68551, 70419, 73014, 75242, 75505, 78067, 78265, 78290, 78956, 79196, 81886, 82399, 83380, 86645, 86704, 88134, 96416, 96429, 96535, 110920, 14117, 117297, 117617, 118870, 123740, 123792, 123876, 123880, 126685, 126696, 126990, 127939, 146146, and 146312.

44. The method of claim 43, wherein the one or more polymorphic variations are in a region spanning positions 78265 to 123880, spanning positions 78265 to 96535, or spanning positions 117297 to 123880 in SEQ ID NO: 1.

45. The method of claim 42, wherein the breast cancer detection procedure is selected from the group consisting of a mammography, an early mammography program, a frequent mammography program, a biopsy procedure, a breast biopsy and biopsy from another tissue, a breast ultrasound and optionally ultrasound analysis of another tissue, breast magnetic resonance imaging (MRI) and optionally MRI analysis of another tissue, electrical impedance (T-scan) analysis of breast and optionally of another tissue, ductal lavage, nuclear medicine analysis (*e.g.*, scintimammography), BRCA1 and/or BRCA2

sequence analysis results, thermal imaging of the breast and optionally of another tissue, and a combination of the foregoing.

46. The method of claim 42, wherein the breast cancer preventative procedure is selected from the group consisting of one or more selective hormone receptor modulators, one or more compositions that prevent production of hormones, one or more hormonal treatments, one or more biologic response modifiers, surgery, and drugs that delay or halt metastasis.

47. The method of claim 46, wherein the selective hormone receptor modulator is selected from the group consisting of tamoxifen, reloxifene, and toremifene; the composition that prevents production of hormones is an aromatase inhibitor selected from the group consisting of exemestane, letrozole, anastrozol, goserelin, and megestrol; the hormonal treatment is selected from the group consisting of goserelin acetate and fulvestrant; the biologic response modifier is an antibody that specifically binds herceptin/HER2; the surgery is selected from the group consisting of lumpectomy and mastectomy; and the drug that delays or halts metastasis is pamidronate disodium.

48. A method of targeting information for preventing or treating breast cancer to a subject in need thereof, which comprises:

detecting the presence or absence of one or more polymorphic variations associated with breast cancer in a nucleic acid sample from a subject, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (b) a nucleotide sequence which encodes a polypeptide encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to the amino acid sequence encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising the polymorphic variation; and

directing information for preventing or treating breast cancer to a subject in need thereof based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

49. The method of claim 48, wherein the one or more polymorphic variations are at one or more positions in SEQ ID NO: 1 is selected from the group consisting of 50479, 50567, 51865, 58349, 58464, 68551, 70419, 73014, 75242, 75505, 78067, 78265, 78290, 78956, 79196, 81886, 82399, 83380,

86645, 86704, 88134, 96416, 96429, 96535, 110920, 14117, 117297, 117617, 118870, 123740, 123792, 123876, 123880, 126685, 126696, 126990, 127939, 146146, and 146312.

50. The method of claim 48, wherein the one or more polymorphic variations are in a region spanning positions 78265 to 123880, spanning positions 78265 to 96535, or spanning positions 117297 to 123880 in SEQ ID NO: 1.

51. The method of claim 48, wherein the information comprises a description of a breast cancer detection procedure, a chemotherapeutic treatment, a surgical treatment, a radiation treatment, a preventative treatment of breast cancer, and combinations of the foregoing.

52. A composition comprising a breast cancer cell and an antibody that specifically binds to a protein, polypeptide or peptide encoded by a nucleotide sequence identical to or 90% or more identical to a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B.

53. A composition comprising a breast cancer cell and a RNA, DNA, PNA or ribozyme molecule comprising a nucleotide sequence identical to or 90% or more identical to a portion of a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B.

54. The composition of claim 53, wherein the RNA molecule is a short inhibitory RNA molecule.

55. A method of selecting a subject that will respond to a treatment of breast cancer, which comprises:

detecting the presence or absence of one or more polymorphic variations associated with breast cancer in a nucleic acid sample from a subject, wherein the polymorphic variation is detected in a nucleotide sequence selected from the group consisting of:

- (a) the nucleotide sequence of Figures 1A-1YYYY or Figures 2A-2B;
- (b) a nucleotide sequence which encodes a polypeptide consisting of an amino acid sequence encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B;
- (c) a nucleotide sequence which encodes a polypeptide that is 90% or more identical to an amino acid sequence encoded by a nucleotide sequence in Figures 1A-1YYYY or Figures 2A-2B; and
- (d) a fragment of a nucleotide sequence of (a), (b), or (c) comprising the polymorphic variation; and

selecting a subject that will respond to the breast cancer treatment based upon the presence or absence of the one or more polymorphic variations in the nucleic acid sample.

56. The method of claim 55, wherein the one or more polymorphic variations are at one or more positions in SEQ ID NO: 1 is selected from the group consisting of 50479, 50567, 51865, 58349, 58464, 68551, 70419, 73014, 75242, 75505, 78067, 78265, 78290, 78956, 79196, 81886, 82399, 83380, 86645, 86704, 88134, 96416, 96429, 96535, 110920, 14117, 117297, 117617, 118870, 123740, 123792, 123876, 123880, 126685, 126696, 126990, 127939, 146146, and 146312.

57. The method of claim 55, wherein the one or more polymorphic variations are in a region spanning positions 78265 to 123880, spanning positions 78265 to 96535, or spanning positions 117297 to 123880 in SEQ ID NO: 1.